DISAGGREGATION TECHNIQUES FOR ESTIMATING RAINFALL OF AN URBAN AREA: A CASE STUDY

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor of Technology
In
Civil Engineering

By

Srijita Jana

Department of Civil Engineering
National Institute of Technology, Rourkela
May, 2011

DISAGGREGATION TECHNIQUES FOR ESTIMATING RAINFALL OF AN URBAN AREA: A CASE STUDY

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor of Technology
In
Civil Engineering

By
Srijita Jana

Under the supreme guidance of Prof. Ramakar Jha

Department of Civil Engineering
National Institute of Technology, Rourkela
May, 2011

National Institute of Technology, Rourkela

CERTIFICATE

This is to certify that the project entitled, "Disaggregation Techniques for Estimating Rainfall for an Urban Area: A Case Study" submitted by 'Srijita Jana' in partial fulfillments for the requirements for the award of Bachelor of Technology Degree in Civil Engineering at National Institute of Technology, Rourkela (Deemed University) is an authentic work carried out by her under my supervision and guidance.

To the best of my knowledge, the matter embodied in the report has not been submitted to any other University/ Institute for the award of any Degree or Diploma.

Date: (Prof. Ramakar Jha)

ACKNOWLEDGEMENT

I express my gratitude and sincere thanks to Prof. Ramakar Jha for his guidance and constant encouragement and support during the course of my work. I truly appreciate the value and his esteemed guidance and encouragement from beginning to the end of the thesis, his knowledge and company at the time of crisis would be remembered lifelong.

My sincere thanks to Prof. M. Panda, Head of the Civil Engineering Department, National Institute of Technology, Rourkela for providing all kinds of help and support. I am also thankful to all the professors of Civil Engineering Department.

I express my special thanks to Prof. R. Mehrotra, University of New South Wales, Australia and Prof. Pradeep Mujumdar, IISc, Bangalore, for extending help when I needed the most. I thank Dr. O.G. B. Sveinsson, Dr. J. D. Salas, W. L. Lane and Dr. D.K Frevert for allowing the free usage of SAMS 2007. I also thank the State Water Data Centre, Gandhinagar for providing the necessary hydrological and climatic data for my research.

I would also like to thank my friends, Sonam Lakra and Mrityunjay Sahu in particular, who have directly and indirectly helped me in the successful completion of my project work.

Lastly I want to thank my parents for constantly supporting and encouraging me at the time of failure and being there beside me always. They have been my pillars of support for these four years. These four years though being separated, has brought me closer to them. I want to dedicate this piece of work of mine to my parents.

SRIJITA JANA
(107CE017)
Department of Civil Engineering,
National Institute of Technology, Rourkela

Table of Contents

ABSTRACT

Disaggregation of rainfall is a very important technique to reduce values from the regional scale to local scale. It is done mainly to study the hydrological changes more effectively over a smaller area at a finer level. In the present study, spatial disaggregation has been carried out for the district of Ahmedabad. The monthly and annual point rainfall values for four stations: Bavla, Nal Lake, Sanand, Warna, have been disaggregated from the mean areal level.

For this purpose, the frequency analysis technique and the disaggregation technique based on Valencia- Schaake Model have been used. In the frequency analysis, it was observed that the R^2 values between the generated and the historical data were as high as 0.98. This indicates the high applicability of the model.

The VS model was used to estimate parameters and generate the point rainfall values for the next 11 years. The results are promising and can be used to estimate more values with approximately 10-15% error.

List of Figures

List of Tables

Chapter I

Introduction

Disaggregation models are a widely accepted tool to generate synthetic hydrologic time series. The limitations and attributes of this model have been reviewed and discussed by several researchers and practitioners. The main objective of this project is to analyze and exploit the potential of disaggregation schemes by steps for Ahmedabad city.

The generation of synthetic series is generally needed for reservoir sizing, for determining the risk of failure (or reliability) of water supply for irrigation system, for determining the risk of failure of dependable capacities of hydroelectric systems and planning studies of future reservoir operation and similar applications (Salas et al., 1980).

The effects of global climate change due to urbanization, on water resources have been studied for several years. It has yet to be explained if these changes are a result of natural variation or greenhouse warming. Global Circulation Models (GCMs) operate with a resolution varying from 2° x 2°to 10° x 10° latitude/longitude. Hence the resulting climatic prediction cannot be directly used for developing localized responses to climate change impacts in the basin/catchment scale using hydrological models. Hydrological processes generally occur at scales in the order of thousands of square kilometers. Also for a better understand of the processes it is important to study the variations in hydrological parameters at local or regional scale (Moss & Tasker, 1987; Riebsame, 1988; Hay et al., 1992; Barros & Lettenmaier, 1993; Bardossy & Plate, 1992).

Very often a hydrological stochastic process must be studied in different time scales. Therefore, the problem arises of how to generate consistent time series both in a coarser, or higher-level, time scale and a finer, or lower-level, time scale. A trivial solution of this problem is to model the process in the lower-level time scale only, and then aggregate to derive the process in the higher-level time scale. However, there are reasons to avoid this solution and model the process in different time scales separately, each time focusing on different important statistical properties of the process. For instance, at the annual scale a model focuses on the long-term persistence properties of the process; at the monthly scale another type of model should describe periodicity and short-term memory of the process; and at the hourly time scale a model should describe intermittency (e.g. wet versus dry state) and fine-scale structure of the process.

In other cases, the higher-level process may be the output of a specialized model (e.g., a meteorological or climatological model) or known from measurements (e.g., daily rainfall measurements); apparently in such cases the aggregation approach cannot work, but rather disaggregation is needed. Specifically when dealing with rainfall, there is a large number of daily raingauges (pluviometres), which have often been operational for a few decades. However, the number of raingauges providing hourly or sub-hourly resolution data (pluviographs, rainfall sensors) is smaller by about an order of magnitude. This situation reflects a general relative paucity of rainfall data for time-scales of one hour or less, both in number of gauges and length of the recorded series. The need for hourly data for hydrological applications, especially in flood studies, suggests the use of appropriate techniques to refine the available daily information and provide the user with possible realizations of hourly precipitation which aggregate up to the given daily data. Such techniques would provide a continuous simulation tool for use for design and management of hydrosystems.

This kind of problems is commonly tackled by disaggregation models. Several such models have been developed since the 1970s and utilized in numerous hydrological applications, including among others, simulation of reservoir systems, either for design or operation purposes, storm and flood simulations, and even enhancement of hydrological data records. Some of the disaggregation models developed are general-purpose as they are not specific to a certain hydrological process or a certain application. General-purpose models, however, are not always applicable to the rainfall process, especially at fine time scales, due to certain peculiarities it exhibits. This triggered the development of specialized techniques for the rainfall process, whose study consisted a preferential field of disaggregation models.

Chapter II

Literature Review

Today, there is renewed interest in disaggregation methods as climate related issues (e.g. downscaling of climate change scenarios) have attracted the interest of many researchers. Usually, such scenarios are developed at a coarse time scale and there is a need to transform them into a finer scale. However, disaggregation is not identical to downscaling, as the latter aims at producing finer scale time series with the required statistics but that do not necessarily add up to any given coarse scale totals. Downscaling is in particular used for hydrological applications of general circulation models (GCM) output where the exact values of the large-scale GCM totals are not considered particularly reliable. In both problem types, synthetic fine-scale series should reproduce the important statistical features of the related hydrological processes at this time scale.

Disaggregation models were introduced in hydrology by the pioneering work of Valencia and Schaake (1972, 1973). Their multivariate model generates sequences of lower-level values at many locations as linear combinations of the related higher-level values and independently generated random components.

The parameter estimation procedure of this model ensures the resemblance of variance and covariance properties between historical and generated series. However the model makes no effort to preserve covariances of the lower level variables belonging to consecutive periods. In fact, this model does not assume any connection between lower-level variables of different periods.

Different model structures and parameter estimation procedures intended to preserve the lagged covariance properties among lower-level variables belonging to consecutive periods have been suggested by Mejia and Rousselle (1976), Hoshi and Burges (1979) and Stedinger and Vogel

(1984). However, as showed by Stedinger and Vogel (1984), the reproduction of the exact historical serial correlations of the lower-level variables is an impossible task within the framework of these models because of structural constraints imposed by that framework, thus giving rise to an inconsistency.

In cases where all variables are normally distributed, the basic linear generating scheme of the Valencia-Schaake model ensures the preservation of higher order moments and distribution functions as well; otherwise the model in its primary formulation fails to maintain such properties. Most of the disaggregation models have been used to disaggregate annual rainfall to monthly amounts. However, as first pointed out by Valencia and Schaake (1972), most modeling schemes of this kind are not suitable for the disaggregation of rainfall for time scales finer than monthly, due to the skewed distributions and the intermittent nature of the rainfall process at fine time scales. Other disaggregation models, different from those described above, have been proposed and used particularly for the disaggregation of rainfall. These, however, do not exhibit the generality of the Valencia-Schaake type linear schemes and in many cases are ad-hoc techniques rather than consistent generalized methods.

All fine time scale rainfall disaggregation techniques summarized above have a common characteristic: they are single-site. The problem of multiple site rainfall disaggregation, as a means for simultaneous spatial and temporal disaggregation, is of significant practical interest but presents significant differences from that of single-site disaggregation, including increased mathematical complexity. The spatial correlation (cross-correlation among different sites) must be maintained in the multivariate problem, whereas it does not appear at all in univariate problems. A first attempt to incorporate more than one site into rainfall disaggregation was done by Socolofsky and Adams (2001) who disaggregated daily rainfall to hourly increments simulating the hourly rainfall trace for each storm from selected intensity patterns measured at a nearby station. In 1998, R Mehrotra et al. carried out spatial disaggregation of rainfall in the Damanganga and Sher Basins. More recently, Kottegoda et al. (2003) simulated daily rainfall through a two-station model at a key station and a satellite station whilst maintaining the first three moments and relevant correlations. Based on this two-station model they were able to perform disaggregation of daily rainfalls into hourly values through dimensionless accumulated hourly amounts generated by a beta distribution, also postulating that the occurrence process of hourly rainfall has a geometric distribution conditioned on the total daily rainfall. Application was made to the Tiber river basin in central Italy. A true multidimensional approach was developed by Koutsoyiannis et al. (2001, 2003) and applied initially to the Brue catchment in South-Western England and subsequently (Fytilas, 2002) to the Tiber river basin in central Italy. Recently a lot of work is being done in mixing these models with Remote Sensing and GIS in more effective studies of river basins.

<u>Chapter III</u>

Study Area

3.1 The District

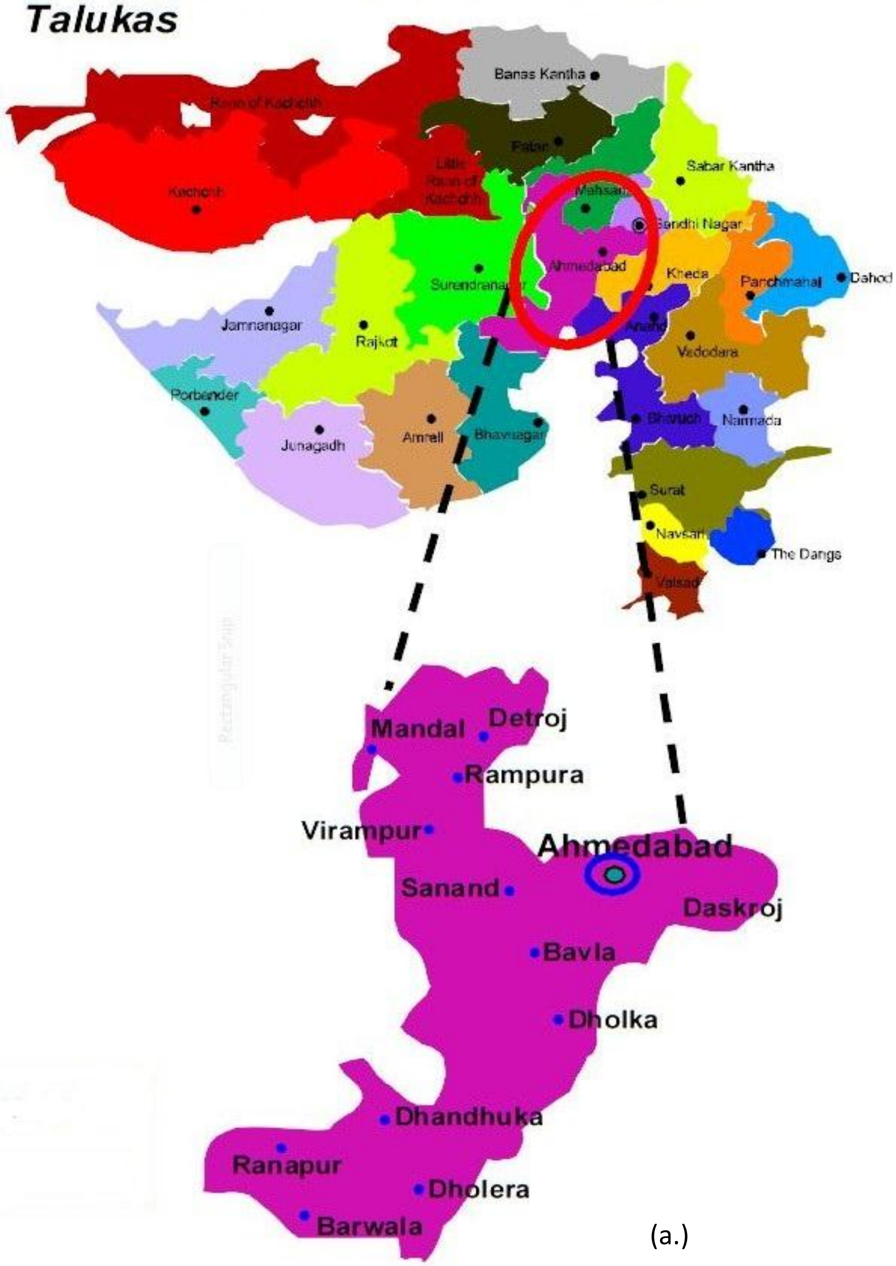
Ahmedabad district is located in the central Gujarat. The district headquarter, Ahmedabad is the largest city in Gujarat and seventh largest urban agglomeration in India. Ahmedabad is spread across ten talukas - Barwala, Daskroi, Dholka, Dhandhuka, Detroj, Sanand, Bavla, Ranpur, Mandal and Viramgam. Ahmedabad is an industrial hub for textiles and is popularly known as the 'Manchester of India'. Due to the presence of several educational institutions of the State, Ahmedabad – Gandhinagar corridor has emerged as an innovative technological and R&D hub. Ahmedabad is developing excellent urban infrastructure for services economy which is largely an urban phenomenon. Key raw materials such as castor, cotton, cumin, fennel, isabgul, potato, are abundantly available.

The city is devoid of any major physical features except for the river Sabarmati, which is cutting the city into two parts: eastern walled city and western Ahmedabad on either side of its banks.

The Ahmedabad-Mumbai Golden Corridor has long been recognized as an important development axis in western India. The city acts as a terminal, rather than as an intermediate node in this linear influence. It has seven major roadways, one expressway and five rail networks. A new corridor between Ahmedabad and Pune has recently emerged, connecting the city to other metropolitan cities including Vadodara, Surat and Mumbai. All these factors have resulted in the axial growth of the region.

Ahmedabad has a tropical monsoon climate, which is hot and dry, except in the rainy season. Summer days are very hot with mean maximum temperature of 41.30C while, nights are pleasant with mean minimum temperature of 26.30C. The mean maximum and minimum temperatures in winter are 300C and 15.40C respectively. The average annual rainfall of the area is 782mm, although there is a considerable variation from year to year. It occurs generally during the months of June to September. The average relative humidity is 60% with a maximum of 80% to 90% during the rainy season.

Map1: District map of Ahmedabad with Talukas



(a.)

(b.)

Figure 1 a.), b.) District Map of Ahmedabad (Source: Internet)

The greater Ahmedabad area has grown at a moderate rate. Growth rates have declined from 3.2 and 2.2 percent (compounded per annum) during the past two decades. However, the rates vary across different spatial units. The population within the AMC limits appears to approach a stabilization level. The areas adjoining AMC, falling within AUDA limits have shown rapid growth. Gandhinagar is also experiencing relatively high rate of growth.

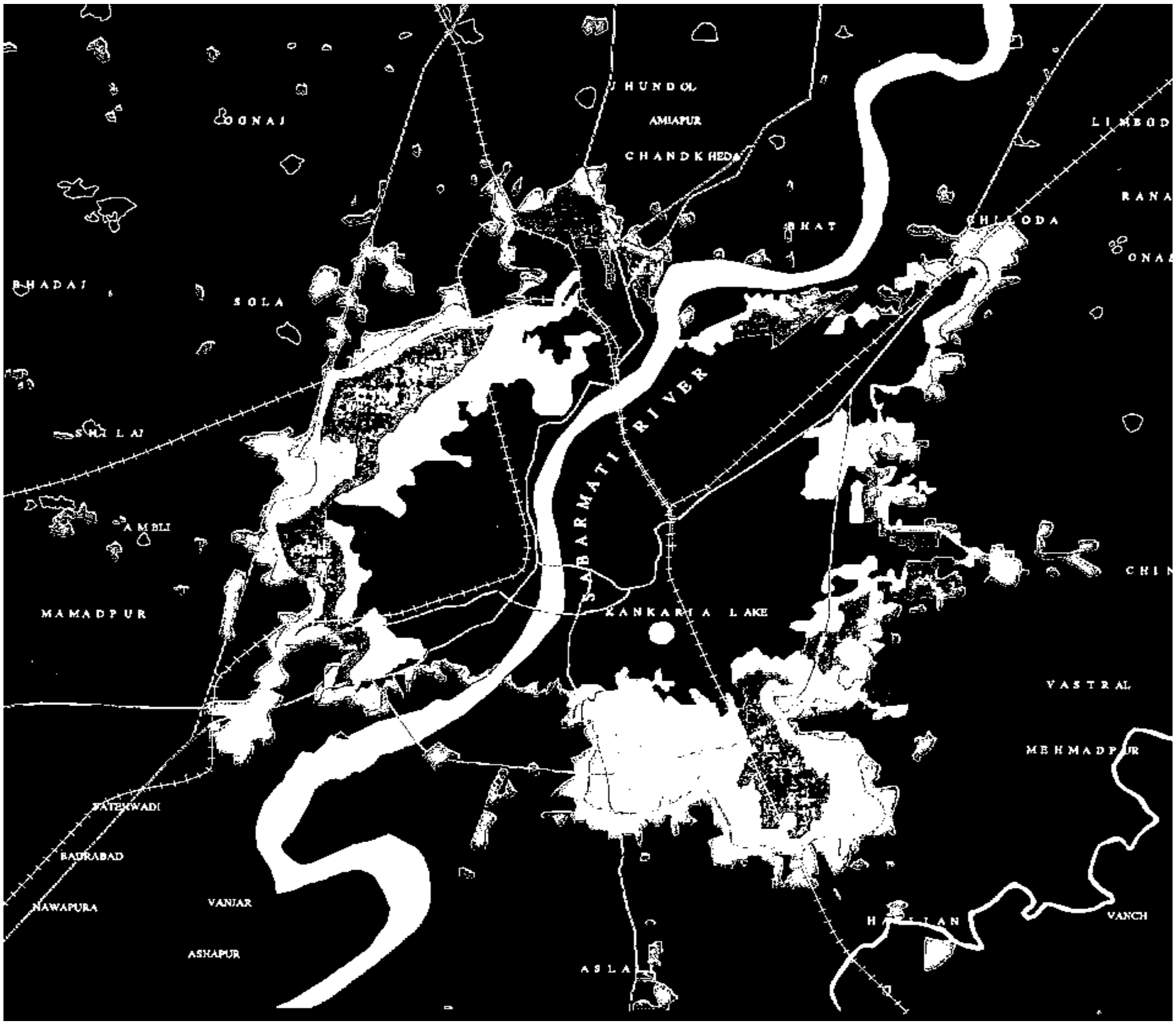


Figure 2: Spatial Growth in the city of Ahmedabad (City Development Plan, Ahmedabad)

The four stations considered in our study are as follows:

Table 1: Geographical Location of Stations

Station	**Latitude**	**Longitude**
Bavla	22°49’40”N	72°22’40” E
Nal Lake	22°49’10” N	72°03’50” E
Sanand	22°59’20” N	72°23’20” E
Warna	22°30’ 30” N	72°24’41” E

Chapter IV

Methodology

4.1 Estimation of Mean Areal Rainfall

The three most widely used methods for the estimation of mean areal rainfall over a catchment area or basin are: i.) Arithmetical- Mean Method ii.) Thiessen Polygon Method, and iii.) Isohyetal Method.

This project makes use of the Thiessen Polygon Method to estimate the mean areal rainfall of the Ahmedabad district. In this method the rainfall recorded at each station is assigned a weightage on the basis of an area closest to the station. The method is as follows. The catchment is drawn to scale and the stations marked on it. Lines are drawn to join each station with the other to form triangles and perpendicular bisectors are drawn for each of the sides of the triangle. These bisectors form a polygon around each station. These bounding polygons are known as Thiessen Polygons and the area of these are determined. If P_1, P_2, P_3,…P_n are the rainfall magnitudes recorded at stations 1, 2, 3..n respectively and A_1, A_2, A_3…A_n are the respective areas of the Thiessen Polygons, the mean areal rainfall of that area shall be:

$$\bar{P}= \frac{\mathrm{P}1A1+P2A2+P3A3+\cdots PnAn}{A1+A2+A3+\cdots An} \tag{4.1}$$

4.2 Cross Correlation of Stations

An important parameter to measure the dependency or the inter-relationship between the stations and the mean areal rainfall is the cross correlation. It is calculated from the following formula.

For multisite series, the lag-k sample cross-correlations between site i and site j, for season τ,

$$\mathrm{r}_k^{ij}= \frac{mkij}{\sqrt{moiimoij}} \tag{4.2}$$

where

$$m_k^{ij} = \frac{1}{N}\sum_{t=1}^{N-k}(yt + k(i) - \bar{y}(i))(yt(j) - \bar{y}(j)) \qquad (4.3)$$

4.3 Auto Correlation of Stations

The sample auto correlation coefficient r_k of a time series is given by: m_k

$$r_{k=} \frac{mk}{mo} \qquad (4.4)$$

where

$$m_k = \frac{1}{N}\sum_{t=1}^{N-k}(yt + k - \bar{y})(yt - \bar{y}) \qquad (4.5)$$

4.4 Frequency Analysis Approach

Frequency analysis is the most common and extensively used statistical approach to deal with the frequency of occurrence of a hydrological variable. The present study considers normal distribution. The Box-Cox power transformation technique for normalizing a series was also used in the analysis.

For each distribution, reduced variates corresponding to values of rainfall at a point (station) were computed using the recommended plotting position formula for that distribution. A simple linear regression equation,

$$Y_t = M_t Z_t + C_t \qquad (4.6)$$

was fitted between the variates (monthly rainfall) *(Y_t)* and the corresponding reduce

using the least square approach and values of the parameters *M*, and *C*, were calculated for station *t*. The same procedure was repeated for all stations and also for mean areal values. Thus if

NS stations are considered, then *NS* + 1 relationships had to be developed and ***NS*** + 1 sets of values of the parameters *M*, and *C*, computed. The relationships could then be used to disaggregate the mean areal rainfall values (Y_{NS+l}) into point values, as follows.

From the established relationship between mean areal values and their reduced variates, $Y_{NS+l} = M_{m+l}Z_{NS+l} + C_{m+l}$, the value of the reduced variate (Z_{NS+i}) corresponding to a mean areal value (Y_{NS+l}) is obtained by substituting the values of parameters M_{m+l} and C_{NS+l} in the above equation. Corresponding to this Z_{NS+1}, the rainfall value *(Y)* at this station is calculated from the developed relationship of station *t* by using the equation $Y_t = M_t Z_{NS+l} + C_t$. Here, Z_{NS+l} is the value of the reduced variate which will not change with station and M_t and C_t, are the parameters for station *t*, already computed.

In the case of a power transformation, the Box-Cox transformation technique was used. Mathematically, the Box-Cox transformation can be represented as:

$$Y_t = (Y_t^a - l)/a \text{ if } a \neq 0 \tag{4.7}$$

$$Y_t = \ln(Y_t) \text{ if } a = 0 \tag{4.8}$$

Here, *Y,* is the monthly rainfall value at a station *t* and value of a is selected in a manner such that the transformed series is normalized and that means that the coefficient of skewness of the transformed series is close to zero.

4.4.1 Error Estimates

Error estimation is done between the generated station values and the historical values. Error estimates are found by subtracting the series value by the mean of the normalized series

$$\varepsilon_i = X_i - \mu, \tag{4.9}$$

and the residual by subtracting the series values by sample mean

$$\widehat{\varepsilon}_i = X_i - \overline{X}. \tag{4.10}$$

The sum of squares of the statistical errors, divided by σ^2, has a chi-square distribution with n degrees of freedom:

$$\sum_{i=1}^{n} (X_i - \mu)^2 / \sigma^2 \sim \chi^2_n. \tag{4.11}$$

The sum of squares of the residuals is observable. The quotient of that sum by σ^2 has a chi-square distribution with only $n - 1$ degrees of freedom:

$$\sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2 / \sigma^2 \sim \chi^2_{n-1}. \tag{4.12}$$

4.5 Spatial Disaggregation by VS Method

For spatial disaggregation of annual data from N key stations to M sub stations there are two models available, namely the Valencia and Schaake (VS) model (Valencia and Schaake, 1973)

$$Y_v = A\, X_v + B\, \varepsilon_v \tag{4.13}$$

and the Mejia and Rousselle (MR) model (Mejia and Rousselle, 1976)

$$Y_v = A\, X_v + B\, \varepsilon_v + C\, Y_{v-1} \tag{4.14}$$

where X_v is the N × 1 column vector of observations in year ν at the N key sites, Y_v is the corresponding M × 1 column vector at the sub sites, ε_v is the M × 1 column noise vector uncorrelated in space and time with each element distributed as standard normal, and A, B, and C are full M × N, M × M, and M × M parameter matrixes, respectively. The differences between the VS and MR models is that the VS model is designed to preserve the lag 0 correlation coefficient in space between all sub stations through the matrix B, and the lag 0 correlation in space between all sub and key stations through the matrix A. The MR model additionally preserves the lag 1 correlation coefficient in space between all sub stations through the matrix C, i.e. the correlations between current year values with past year values.

4.5.1 Modeling of key station

The univariate ARMA model is used to model the mean areal rainfall values. The ARMA(p,q) model of autoregressive order p and moving average order q is expressed as:

$$Y_t = \sum_{i=1}^{p} \varphi i \, Yt - i + \varepsilon_t - \sum_{j=1}^{q} \theta j \, \varepsilon t - j \tag{4.15}$$

where Y_t represents the rainfall process for year t, it is normally distributed with mean zero and variance $\sigma^2(Y)$, ε_t is the uncorrelated normally distributed noise term with mean zero and variance $\sigma^2(\varepsilon)$, $\{\varphi 1,\dots,\varphi p\}$ are the autoregressive parameters and $\{\theta 1,\dots, \theta q\}$ are the moving average parameters. The values of p and q have been taken as 2 each with the method of moments for estimation of parameters.

4.5.2 Parameter Estimation

The model parameter matrixes A and B of the VS model can be estimated by using Method of moments (Valencia and Schaake, 1973):

$$A = M_o(YX)M^{-1}(X) \tag{4.16}$$

$$BB^T = M_o(Y) - AM_o(X)A^{-1} \tag{4.17}$$

$G = B\,B^T$ is the noise variance-covariance matrix (B is the Cholesky decomposition of G), and

$$M_k(Y) = E[\, Y_v Y_{v-k}^T] \text{ and } M_k(YX) = E[\, Y_v X_{v-k}^T]. \tag{4.18}$$

Chapter V

Results

5.1 Mean Areal Rainfall

Using the equation 4.1, the Thiessen Polygons were found out for the four stations (Refer Fig. 3) and the following table, Table 2, is the weight calculated for each.

Table 2: Weights of Stations

Sr. No.	Station Name	Weight
1.	Bavla	0.147
2.	Nal Lake	0.077
3.	Sanand	0.199
4.	Warna	0.578

These weights are then used to calculate the mean monthly and annual areal rainfall of the Ahmedabad district. Table 3 gives the values of the annual rainfall values at each of the stations from 2001-2009 and the areal rainfall calculated based on the weights obtained.

Table 3: Annual Rainfall of stations and mean areal rainfall

Year	Bavla	Nal Lake	Sanand	Warna	Areal
2001	1123.5	952	530	521	645.0665
2002	272	183.5	319.4	184	224.0261
2003	726	627.5	739.4	701	707.3851
2004	780.5	826	1051.6	596	732.0919
2005	1195	1120	1073	432	725.128
2006	968	544	878.5	252	504.6615
2007	550	814.5	913.5	312	505.689
2008	426	973.5	860.5	516	607.069
2009	250	289.5	363.5	220	251.923

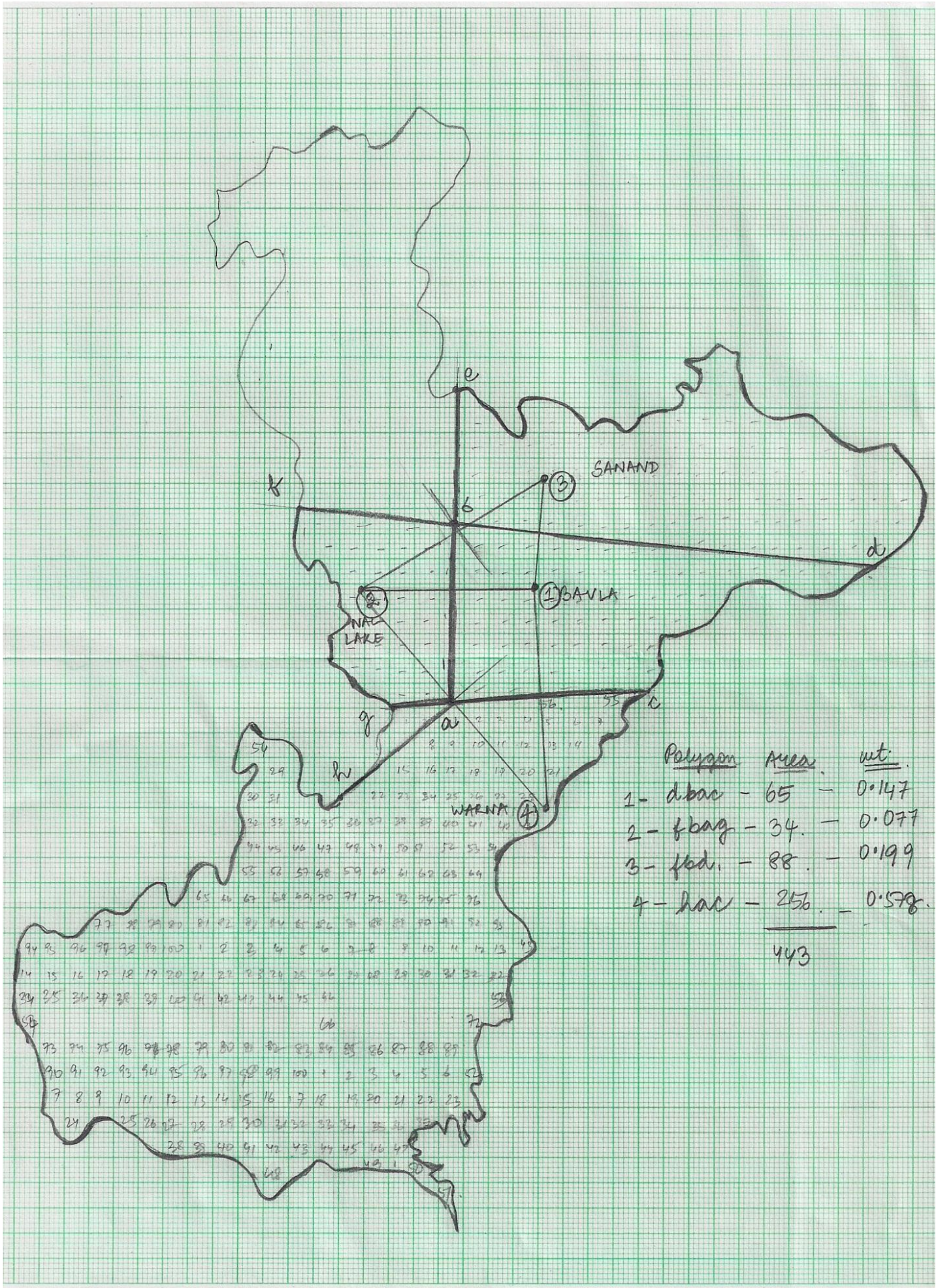


Figure 3: Thiessen Polygon for Ahmedabad District

Figure 4 is a bar chart representation of the annual rainfall values at each of the four stations.

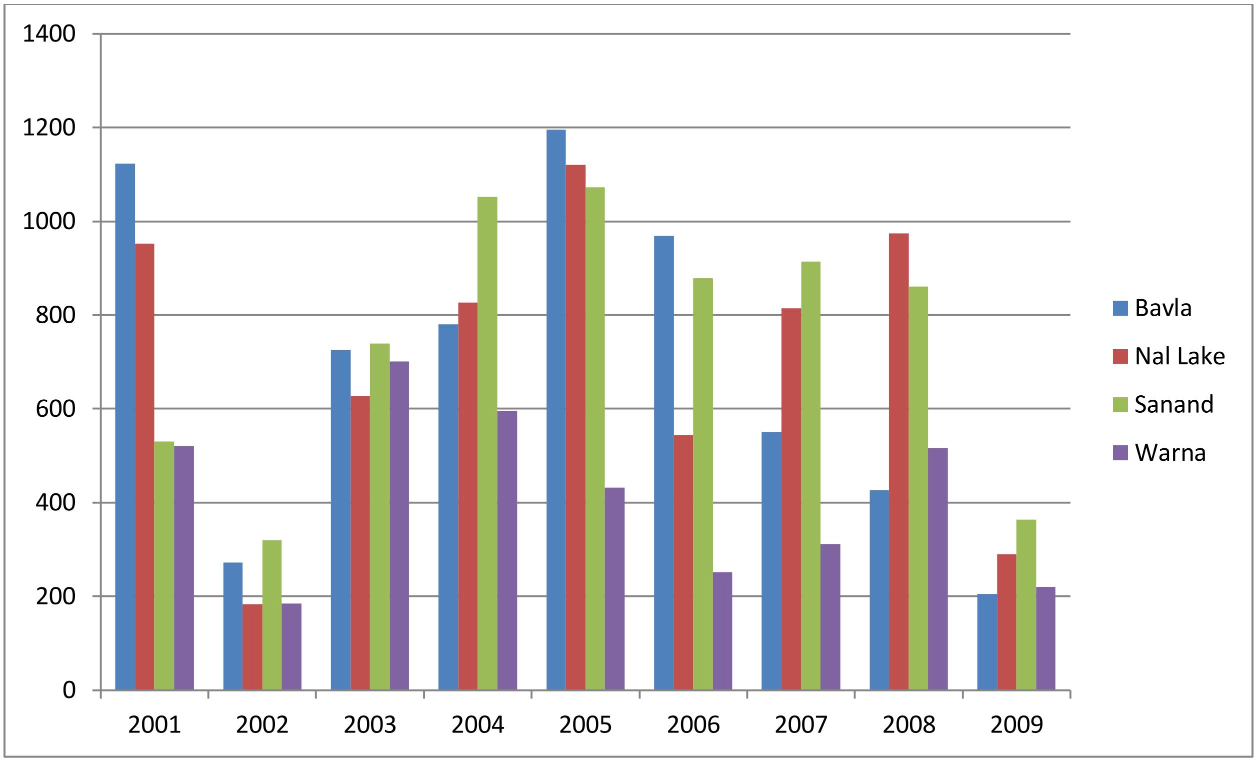


Figure 4: Bar Chart showing Annual Rainfall at four stations

5.2 Cross Correlation of Rainfall with Mean Areal Rainfall

Once the mean annual areal rainfall has been found, the cross correlation between the stations is plotted against time lag to find out how and which station can be substituted in case of failure of instruments at any particular station. A maximum of 15 was taken as the time lag and substituted in Eq. 4.2 and 4.3.

The correlation matrix for Station Bavla and mean areal rainfall is:

$$\begin{bmatrix} 1.000 & 0.7407 \\ 0.7407 & 1.000 \end{bmatrix}$$

The correlation matrices of other stations are given in Appendix A1.

Figure 5 shows the cross correlation between Station Bavla and the mean annual areal rainfall. It is seen here that at lag 0 the values have the highest correlation of around 0.75. The cross correlation plots of other stations with mean areal value is provided in Appendix A2. It was observed that the other stations too displayed fairly good correlations with the mean areal rainfall.

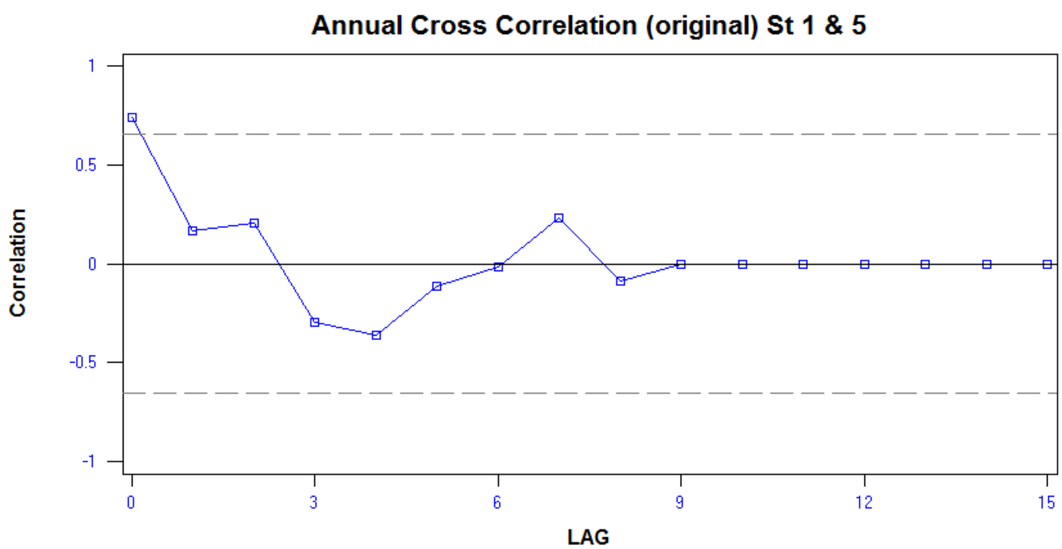


Figure 5: Plot of Annual Correlation between Station
Bavla and Mean Areal Rainfall

5.3 Auto Correlation Factor (ACF) of Rainfall

Next, the auto correlation factor is found to develop a statistical model. This is done to see how the previous period data is related to the present data and how the present data will be related to the generated data.

The auto correlation factors were plotted for all the four stations against time-lag (a maximum of 15) using Eq. 4.4 and 4.5. Figure 6 shows the plot for Station Bavla. For all other stations please refer Appendix A3.

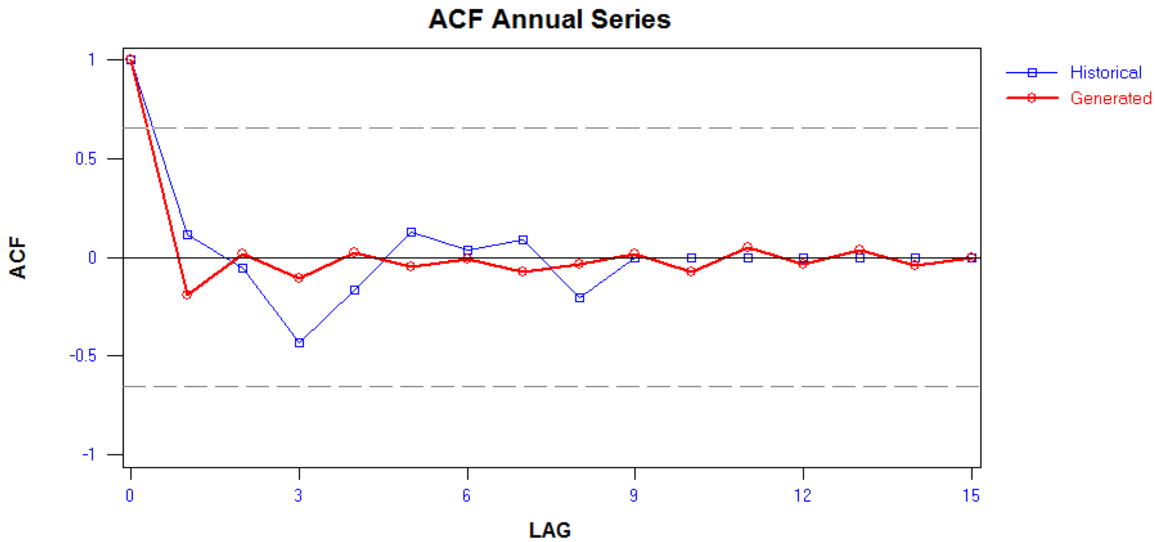


Figure 6: Plot of Auto Correlation Factor (ACF) against time lag for Station Bavla

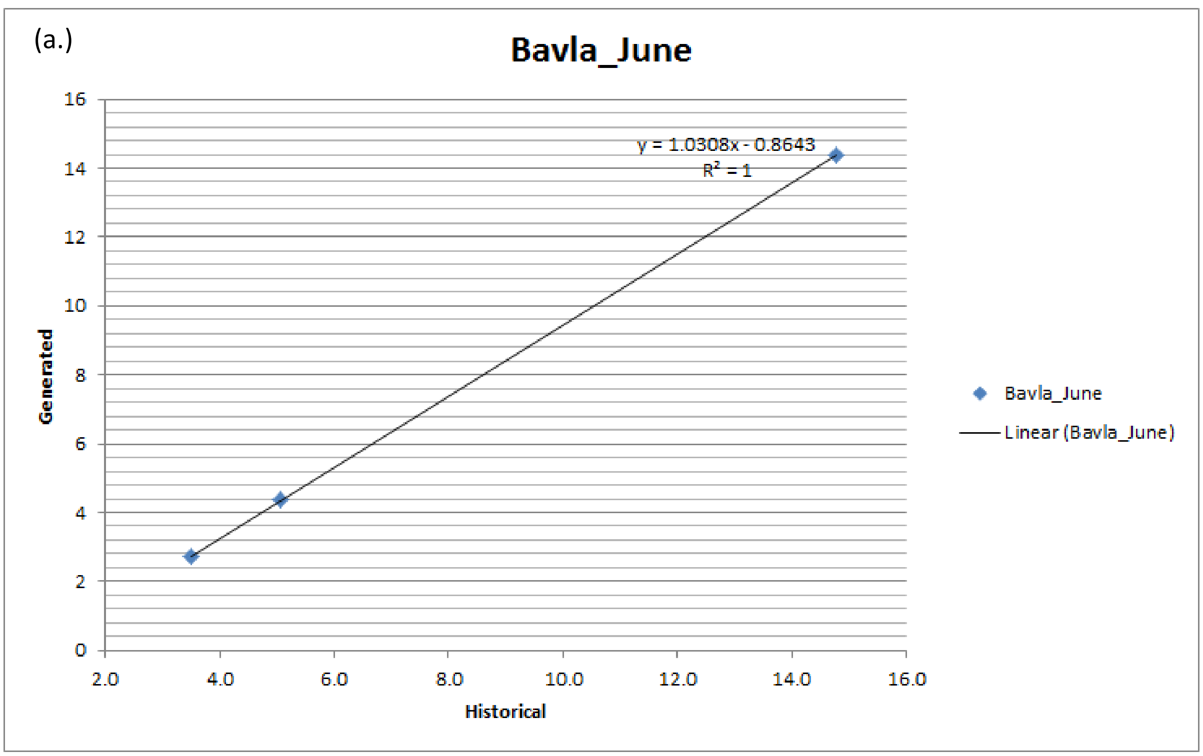
5.4 Frequency Analysis Approach

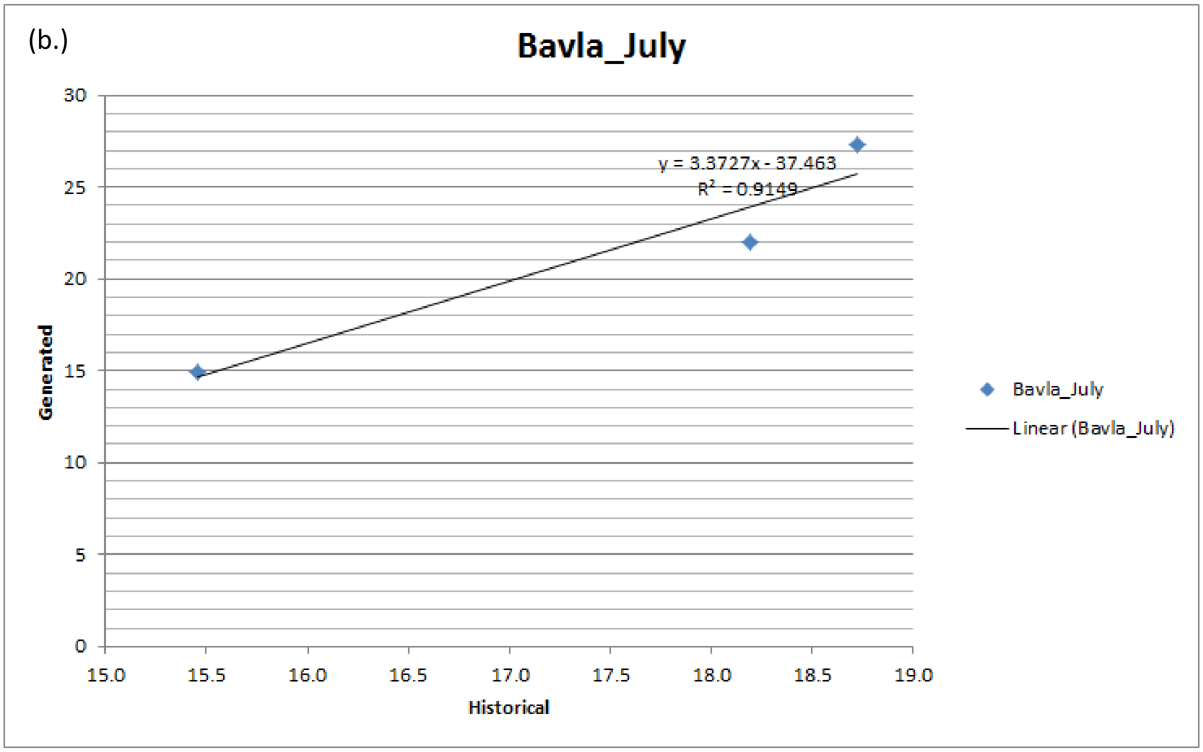
After ensuring proper correlation between the stations and mean areal rainfall values frequency analysis is carried out for each station. In this approach the monthly values for each station and the mean areal rainfall values are first normalized using the Box-Cox Transformation (refer Eq. 4.7 and 4.8). These transformed values were then fitted into a general regression model using Eq. 4.6. Least Square Method was used for the estimation of the parameters. For calibration of the model the monthly values for the period 2001-2006 were used and the values of M and C estimated for each month (June- September). Table 4 below shows the results obtained for Station Bavla. The procedure is repeated for all four stations. The value of parameters for the other three stations, Nal Lake, Sanand and Warna, can be found in Appendix A4.

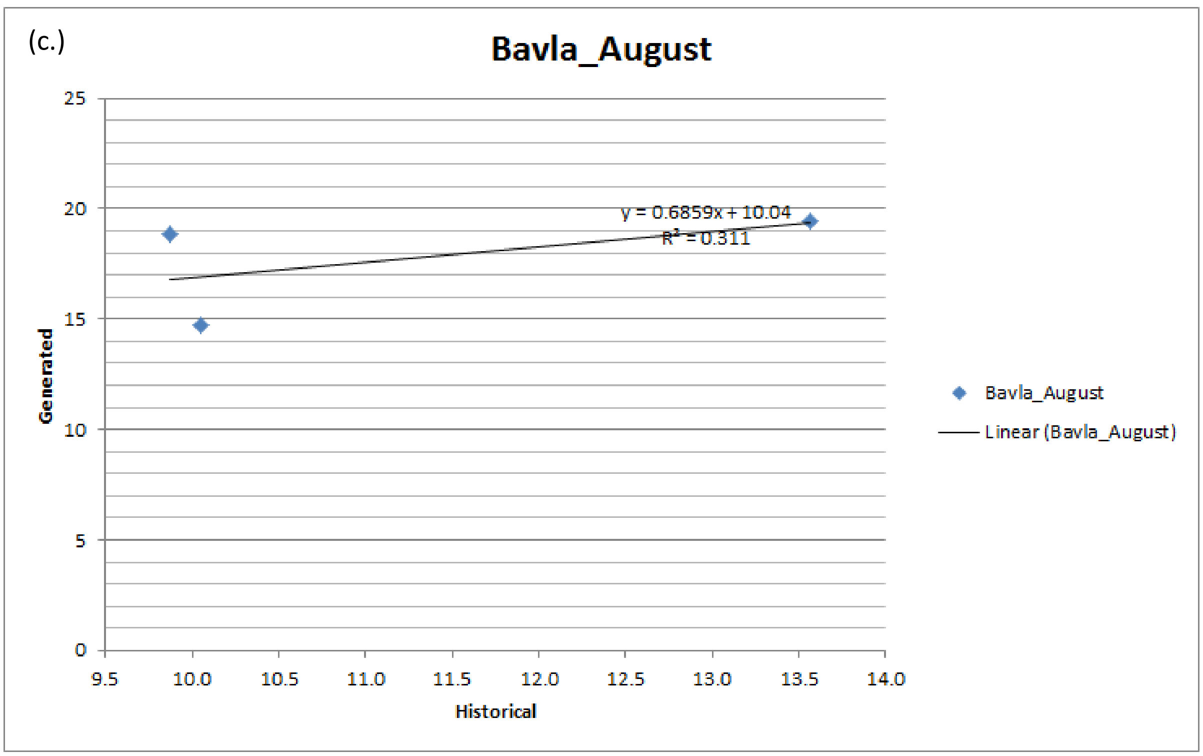
Table 4: Parameter Estimates for Station Bavla using reduced variates by least square method

Month	M	C	R^2
June	0.8044	-0.1623	0.8564
July	0.89953	-3.92902	0.871662
August	0.435191	7.237769	0.681926
September	0.573873	2.778662	0.978408

The regression equation thus obtained is then used to estimate the monthly values for the period 2007-2009 using the mean areal rainfall values during the same period. The generated values were then plotted against the historical values and the residuals obtained for each station. Figure 7a.)- d.) below show the plots between the historical and the generated values for Station Bavla. For all other stations please refer Appendix A5.




(b.)
Bavla_July
y = 3.3727x - 37.463
R² = 0.9149
Generated
Historical
Bavla_July
Linear (Bavla_July)


(c.)
Bavla_August
y = 0.6859x + 10.04
R² = 0.311
Generated
Historical
Bavla_August
Linear (Bavla_August)

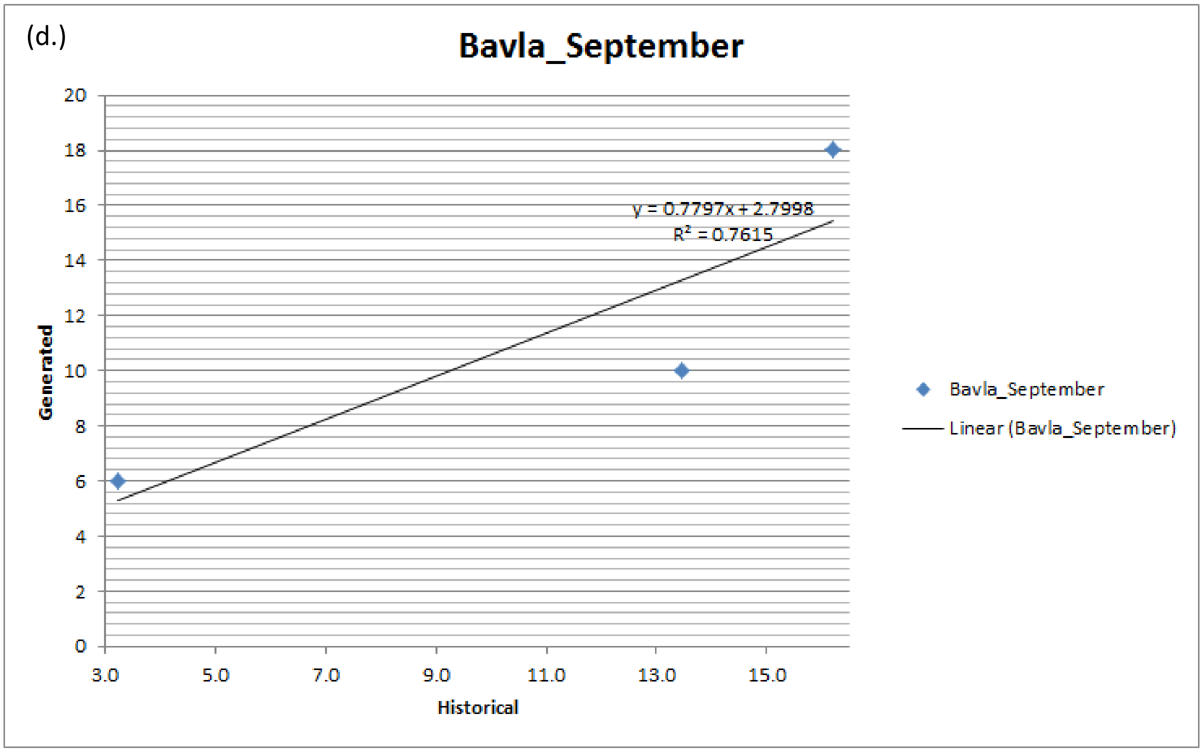


Figure 6: Plot of Historical Vs Generated data for Station Bavla for the month of a.)June b.)July c.)August d.)September

It can be seen from the tables that the correlation is quite high between the regional and the local values. The generated values showed high correlation coefficient with the historical values. It was as high as 0.98 for one case, but even went as low as 0.22 for another.

5.5 Disaggregation by VS Model

For the Valencia- Schaake Model, Eq 4.13 was used to model a relationship between the annual rainfall values of individual stations and the mean areal rainfall. For this approach the mean areal rainfall was taken as key station values and modeled using ARMA model. An autoregressive order of 2 and moving average order of 2 is used to create the model. Eq. 4.15 is used to model the station. The various parameters determined from this model: statistical, AR, MA are given below in Table 5. Parameter estimation in VS Method is done by Method of Moments as discussed in 4.5.2.

5.5.1 Model Parameters

Table 5: Parameters of generated model

Statistical Parameters	Mean	544.79
	Variance	33307.3
	AICC	137.394
	SIC	116.183
	White Noise	55978.4
AR Parameters	Φ_1	0.169992
	Φ_2	0.571092
MA Parameters	θ_1	-0.133217
	θ_2	0.323799

Moment Matrices

A Matrix

$$\begin{bmatrix} 1.372206 \\ 1.392882 \\ 1.126218 \\ 0.807814 \end{bmatrix}$$

B Matrix

$$\begin{matrix} 227.15 & 0 & 0 & 0 \\ 19.7872 & 157.536 & 0 & 0 \\ -8.51817 & 42.528 & 162.23 & 0 \\ -57.4731 & -35.6285 & -55.8546 & 0.00022048 \end{matrix}$$

G Matrix

$$\begin{matrix} 51597.2 & 4494.66 & -1934.9 & -13055 \\ 4494.66 & 25209.1 & 6531.14 & -6750.01 \\ -1934.9 & 6531.14 & 28199.7 & -10086.9 \\ -13055 & -6750.01 & -10086.9 & 7692.28 \end{matrix}$$

5.5.2 General Statistics of Generated Data

For Station Bavla, Table 6 shows the statistics of the historical and the generated data. For statistical details on other stations please refer Appendix A6.

It can be seen here that data generated has close relation to the historical one. The model is seen to give satisfactory results for the other stations also.

Table 6: Statistics of Station Bavla

Paramter	**Historical**	**Generated**
Mean	694.0	669.7
Std Dev	338.1	323.6
CV	0.4872	0.4828
Skew	0.03426	0.1301
Min	205.0	111.1
Max	1195	1337
ACF(1)	0.1155	-0.1905
ACF(2)	-0.05656	0.01849

5.5.3 Time Series Plot for Station Bavla

The generated values of annual rainfall for each station are then plotted along with the historical data. In this model the annual rainfall is generated for four stations for the period 2001-2009 using the mean areal rainfall value (Column 6, Table 3).

The time series between the historical and generated data for Station Bavla is given in Figure 8. For other stations kindly refer Appendix A7.

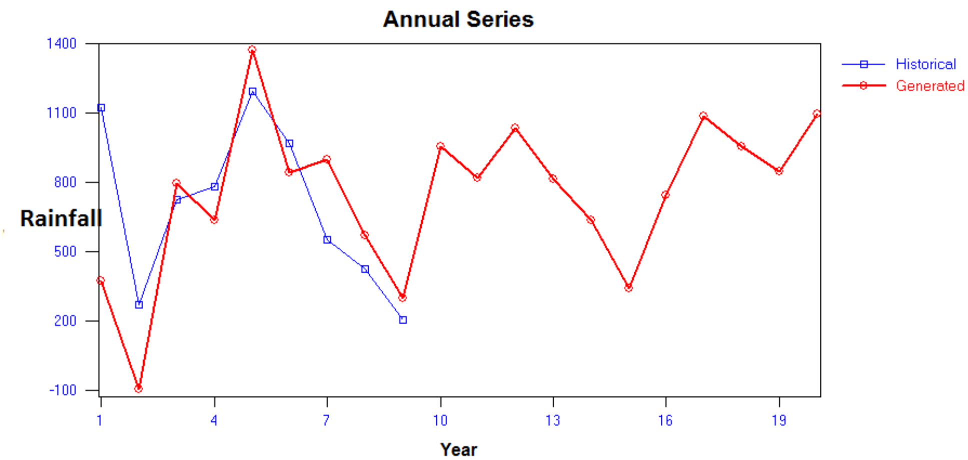


Figure 8: Plot of Historical Vs Generated data for Station Bavla

5.6 Conclusions

i.) It was found that disaggregation technique is very vital to downscale regression data to local scale for higher scale model, for the effective analysis of hydrological processes.

ii.) Results obtained using frequency analysis technique was tested for their validity and were found to give high correlation coefficient.

iii.) Disaggregation Method using the Valencia- Schaake Model was used to further generate the rainfall scenario for the four stations of the Ahmedabad district and the results were found to be very promising.

iv.) The disaggregation model may be used to model rainfall data at daily and hourly levels.

References

1. Brockwell, P.J. and Davis, R.A., 1996. Introduction to Time Series and Forecasting. Springer Texts in Statistics. Springer-Verlag, first edition.
2. Demetris Koutsoyiannis, Rainfall Disaggregation Methods: Theory And Applications, Workshop on Statistical and Mathematical Methods for Hydrological Analysis – Rome May 2003
3. Fernandez, B., and J.D. Salas, 1990, Gamma-Autoregressive Models for Stream-Flow Simulation, ASCE Journal of Hydraulic Engineering, vol. 116, no. 11, pp. 1403-1414.
4. Jawaharlal Nehru National Urban Renewal Mission: City Development Plan Ahmedabad, 2006-2012, AMC, AUDA, CEPT University, Ahmedabad
5. Lane, W.L., 1981, Corrected Parameter Estimates for Disaggregation Schemes, Inter. Symp. On Rainfall Runoff Modeling, Mississippi State University.
6. Lane, W.L., and D.K. Frevert, 1990, Applied Stochastic Techniques, personal computer version 5.2, users manual, Bureau of Reclamation, U.S. Dep. of Interior, Denver, Colorado.
7. R. Mehrotra, R. D Singh, 1998, Spatial Disaggregation of Rainfall Data, Hydrological Sciences, February 1998
8. Salas, J.D., Delleur, J.W., Yevjevich, V., and Lane, W.L., 1980. Applied Modeling of Hydrologic Time Series. Water Resources Publications, Littleton, CO, USA, first edition. Fourth printing, 1997.
9. Salas, J.D., 1993. Analysis and Modeling of Hydrologic Time Series, chapter 19. Handbook of Hydrology. McGraw-Hill.

10. Salas, J.D., Saada, N., Chung, C.H., Lane, W.L. and Frevert, D.K., 2007, "Stochastic Analysis, Modeling and Simulation (SAMS) Version 2007 - User's Manual", Colorado State University, Water Resources Hydrologic and Environmental Sciences, Technical Report Number 10, Engineering and Research Center, Colorado State University, Fort Collins, Colorado.
11. Shairul Rohaziawati Samat, 2004, Stochastic Simulation Of Rainfall Time Series Using Disaggregation Models, Universiti Teknologi Malaysia, Masters Thesis, March 2004
12. Sveinsson, O. G. B., Salas, J. D., and D. C. Boes, 2005: Prediction of extreme events in Hydrologic Processes that exhibit abrupt shifting patterns. Journal of Hydrologic Engineering, 10(4):315-326.
13. Valencia, R.D. and Schaake Jr, J.C., 1973. Disaggregation Processes in Stochastic Hydrology. Water Resources Research, 9(3):580-585.

Appendix

A1. Cross Correlation Matrices of Stations

Station Nal Lake and Mean Areal Rainfall

$$\begin{bmatrix} 1.000 & 0.848 \\ 0.848 & 1.000 \end{bmatrix}$$

(a.) Station Sanand and Mean Areal Rainfall

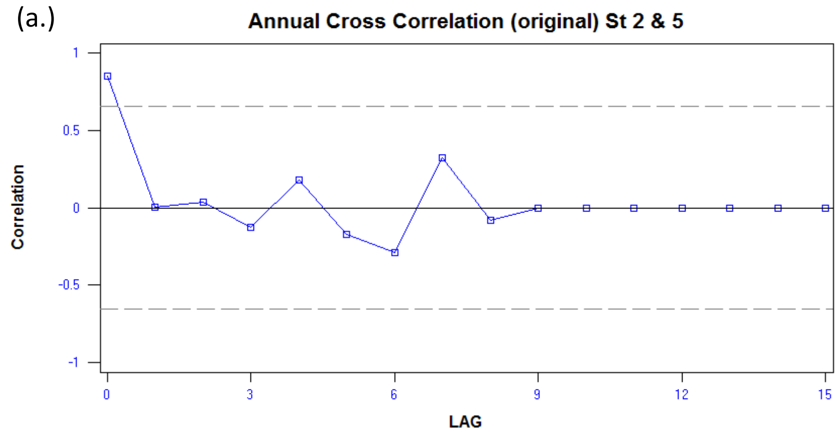
$$\begin{bmatrix} 1.000 & 0.7774 \\ 0.774 & 1.000 \end{bmatrix}$$

(b.) Station Warna and Mean Areal Rainfall

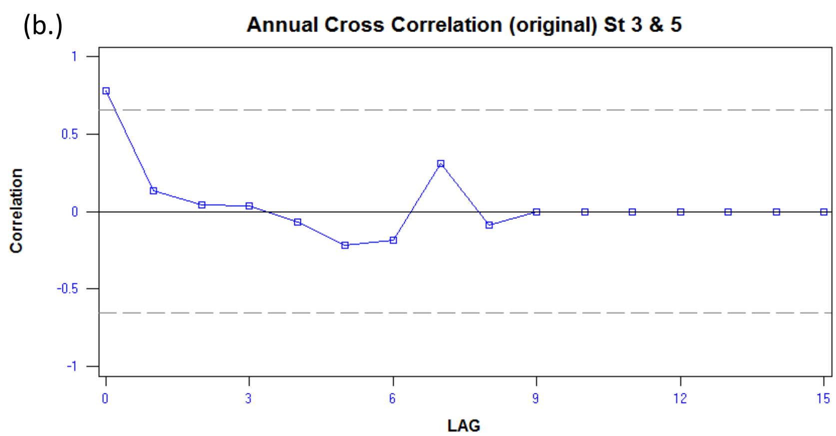
$$\begin{bmatrix} 1.000 & 0.859 \\ 0.859 & 1.000 \end{bmatrix}$$

A2. Cross Correlation Plots of Stations

(a.)



(b.)



(c.)

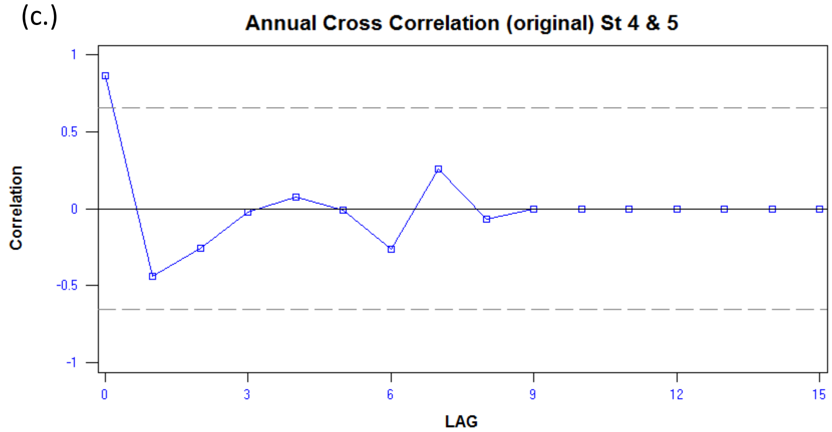
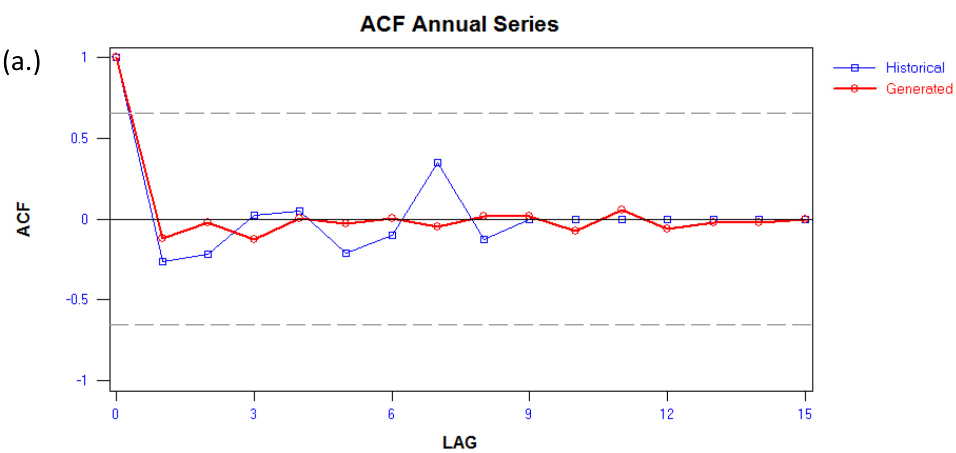


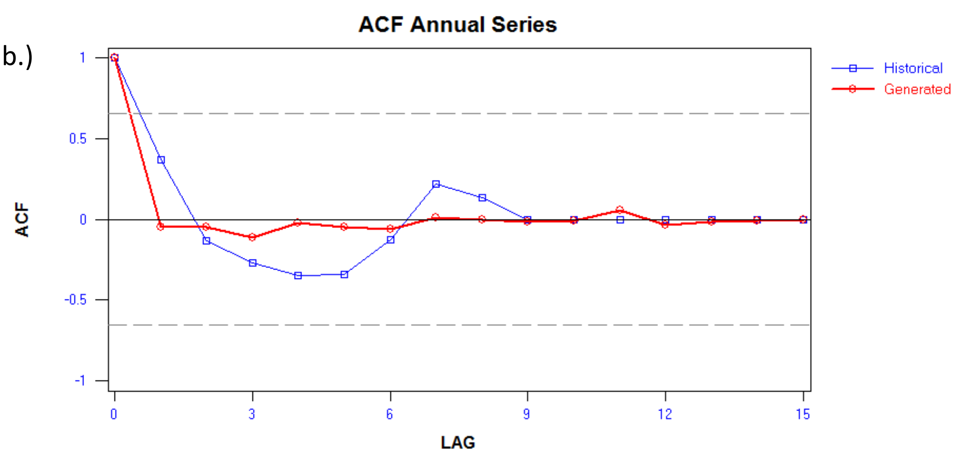
Figure 9: Annual Cross Correlation of Stations a.)Nal Lake b.)Sanand c.)Warna with mean areal rainfall

A3. Auto Correlation Factor

(a.)



(b.)



(c.)

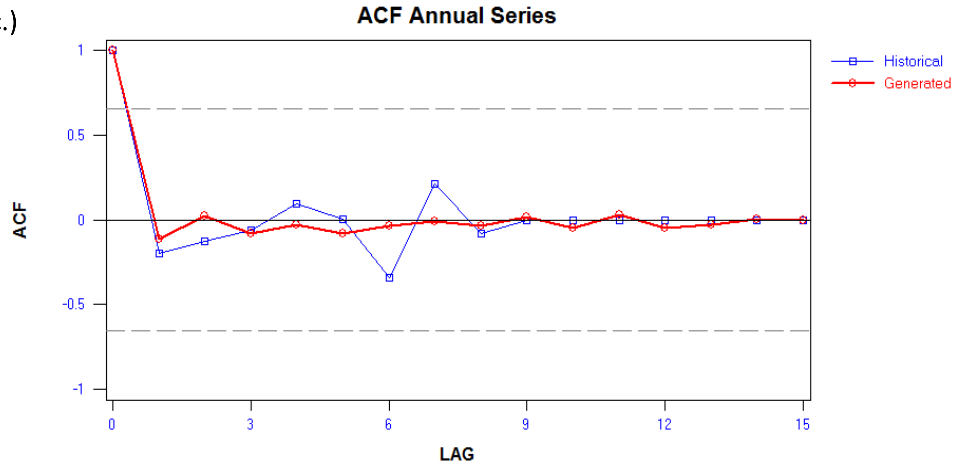


Figure 10: Auto Correlation Factor of Stations a.)Nal Lake b.)Sanand c.)Warna with mean areal rainfall

A4. Frequency Analysis- Calibration of Model

Table 7: Parameter Estimates for Station a.)Nal Lake b.)Sanand c.)Warna using reduced variates by least square method

a.)

Month	**M**	**C**	**R^2**
June	1.1748	2.8102	0.85519
July	0.5333	10.79491	0.0132221
August	0.888586	-0.027104	0.847342
September	0.931740	1.596853	0.478155

b.)

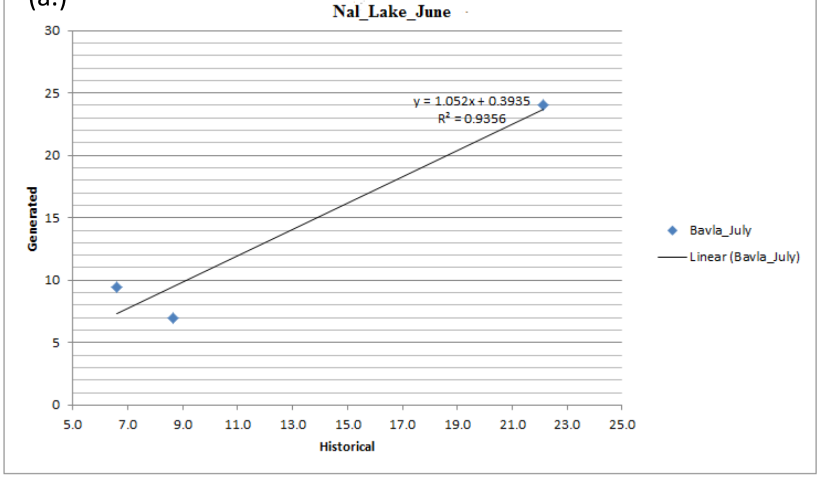
Month	**M**	**C**	**R^2**
June	0.07700	10.53822	0.222532
July	0.43637	12.87006	0.223738
August	0.736059	6.003485	0.910393
September	0.962577	-0.813013	0.890768

c.)

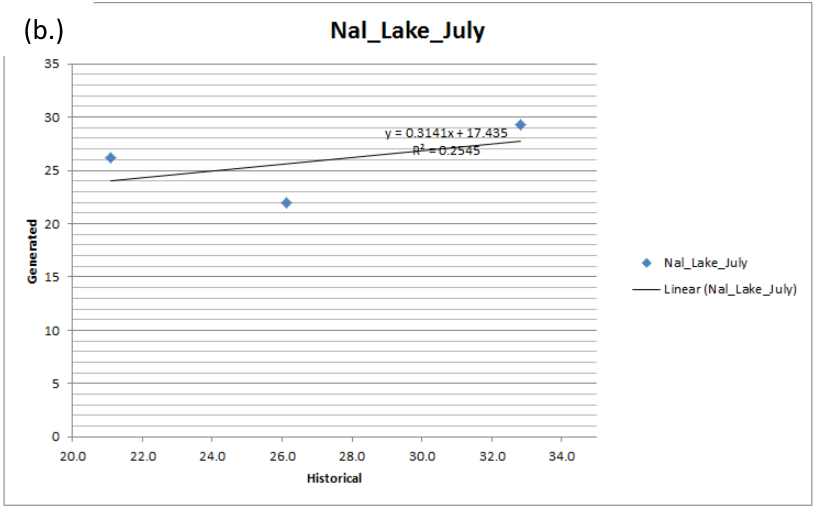
Month	M	C	R^2
June	0.193181	3.514798	0.144321
July	-0.005963	7.963249	-0.332628
August	0.149550	3.791302	-0.600282
September	0.202970	2.416142	0.793242

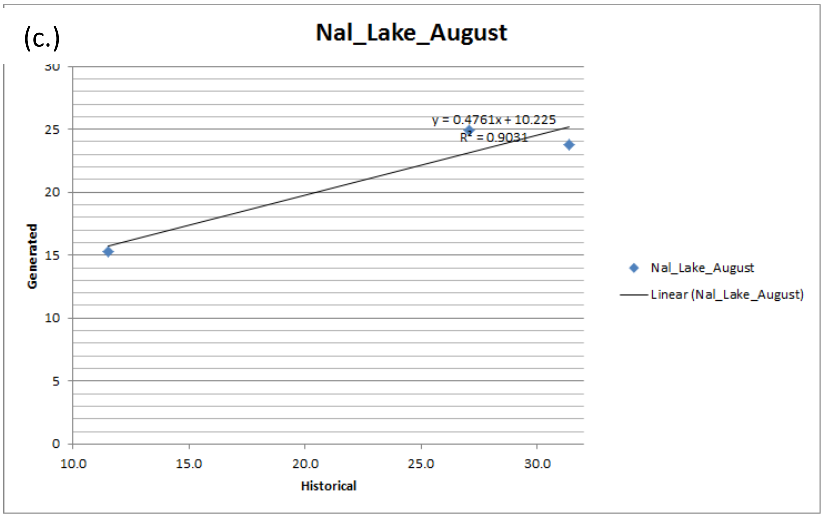
A5. Frequency Analysis- Testing of Model

(a.)



(b.)





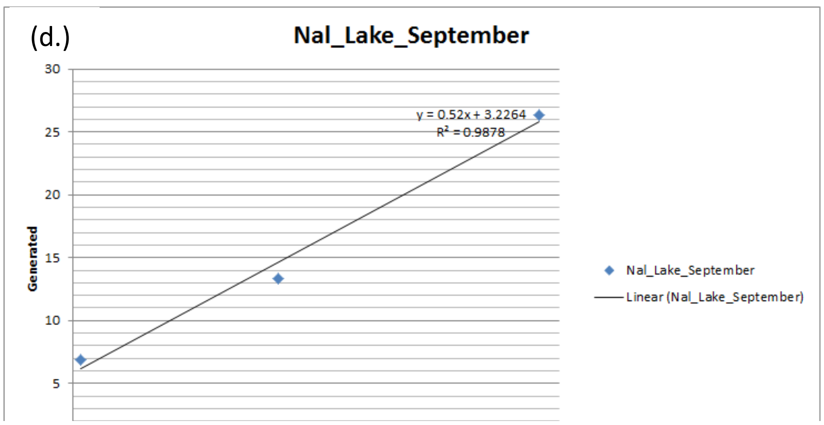
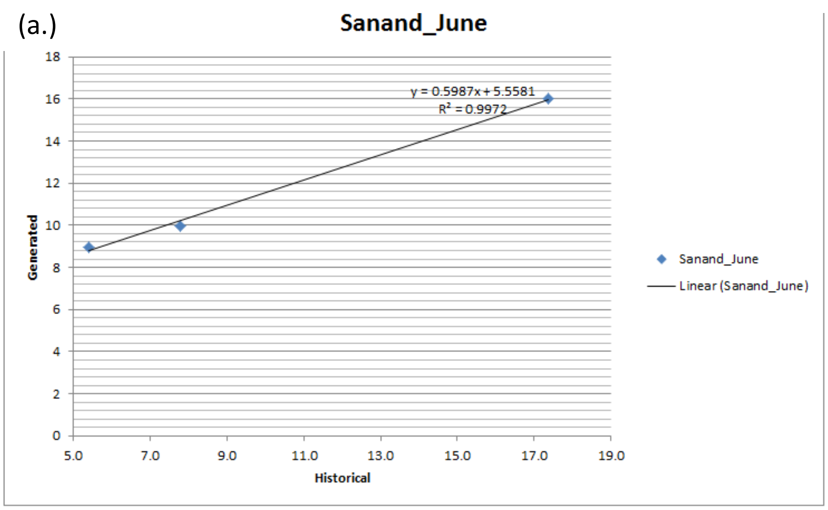


Figure 11: Plot of Historical Vs Generated Data for the station Nal Lake

(a.) June (b.)July (c.)August (d.)September



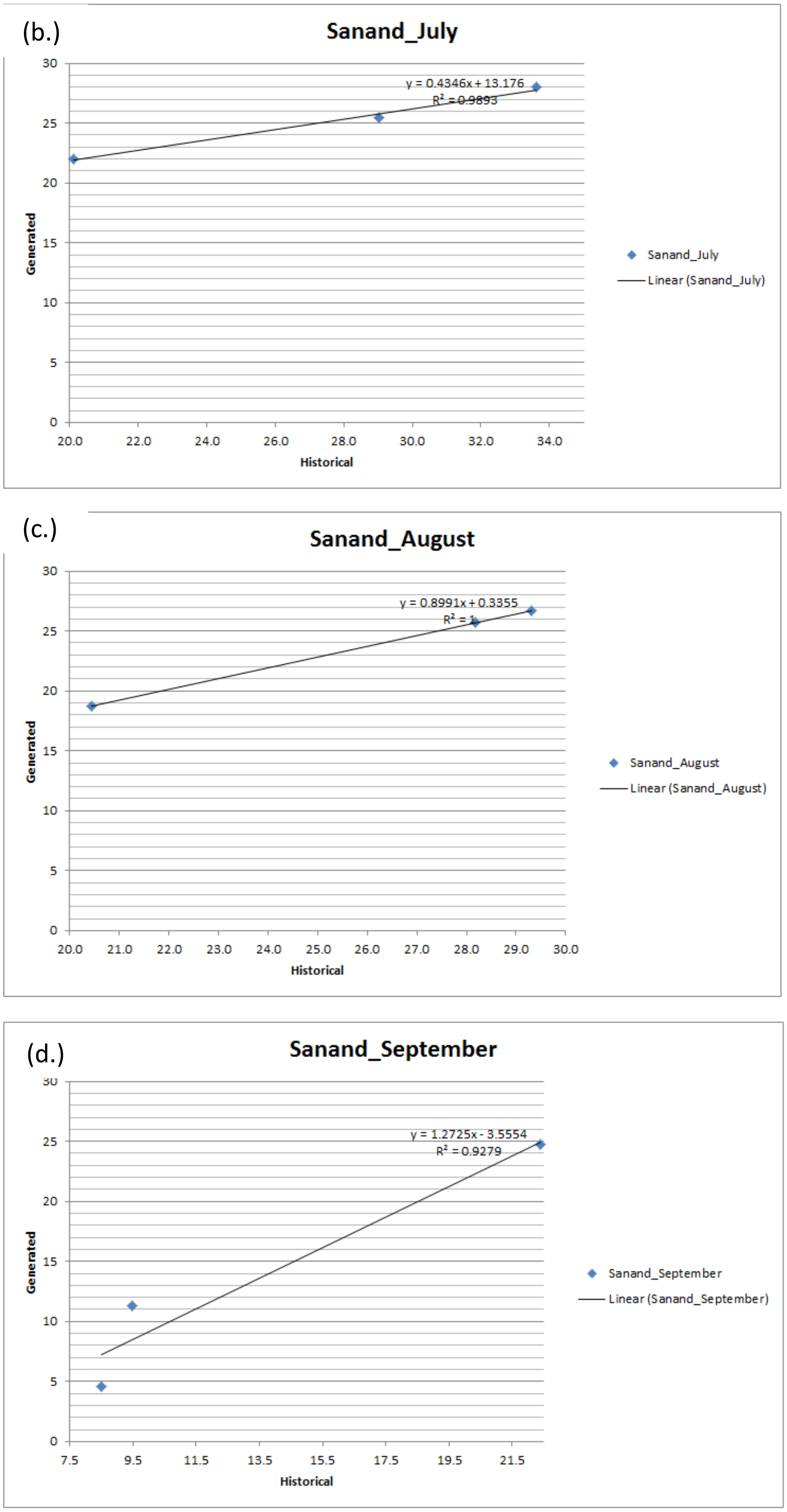
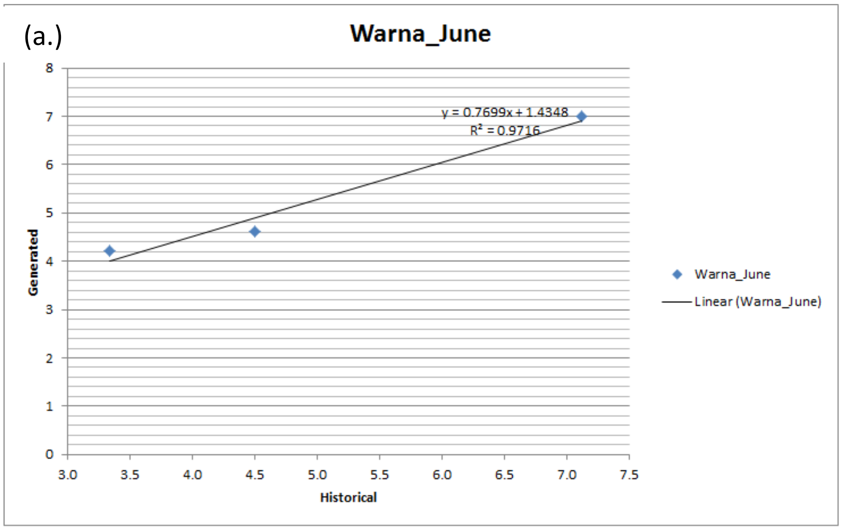


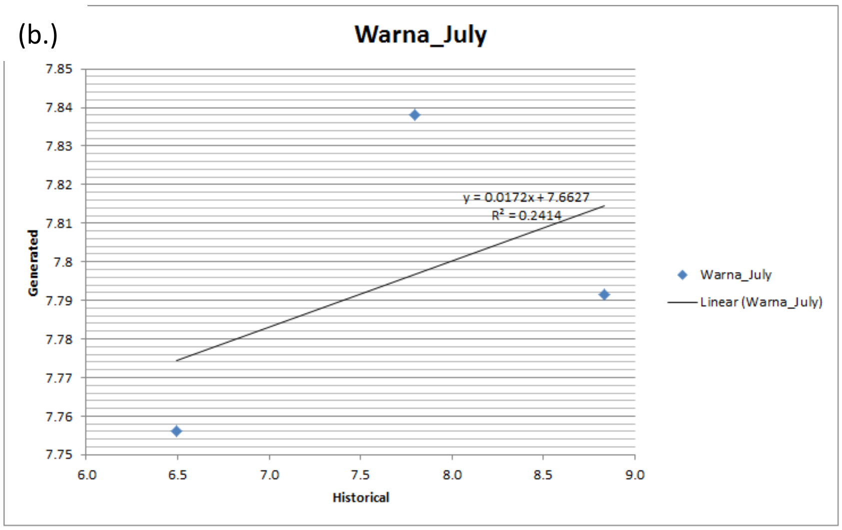
Figure 12: Plot of Historical Vs Generated Data for the station Sanand

(a.)June (b.)July (c.)August (d.)September

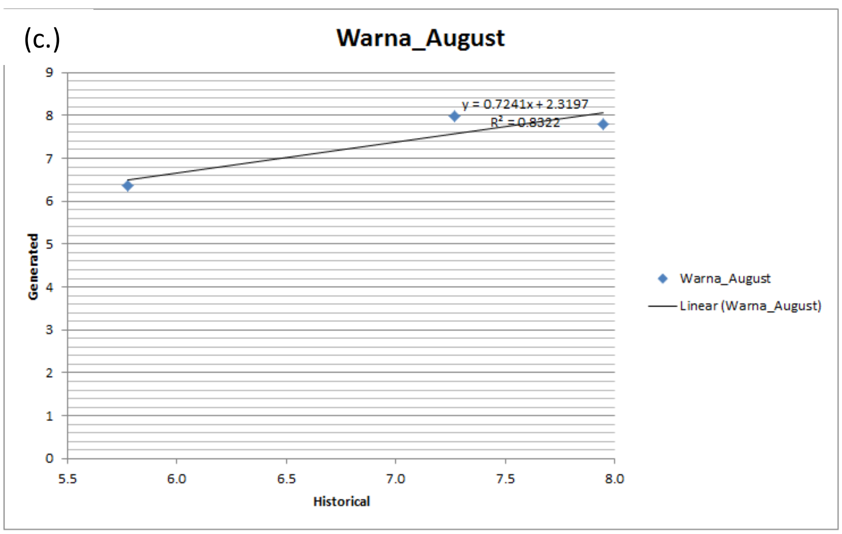
(a.)


Warna_June
y = 0.7699x + 1.4348
R² = 0.9716
Generated
Historical
Warna_June
Linear (Warna_June)

(b.)


Warna_July
y = 0.0172x + 7.6627
R² = 0.2414
Generated
Historical
Warna_July
Linear (Warna_July)

(c.)


Warna_August
y = 0.7241x + 2.3197
R² = 0.8322
Generated
Historical
Warna_August
Linear (Warna_August)

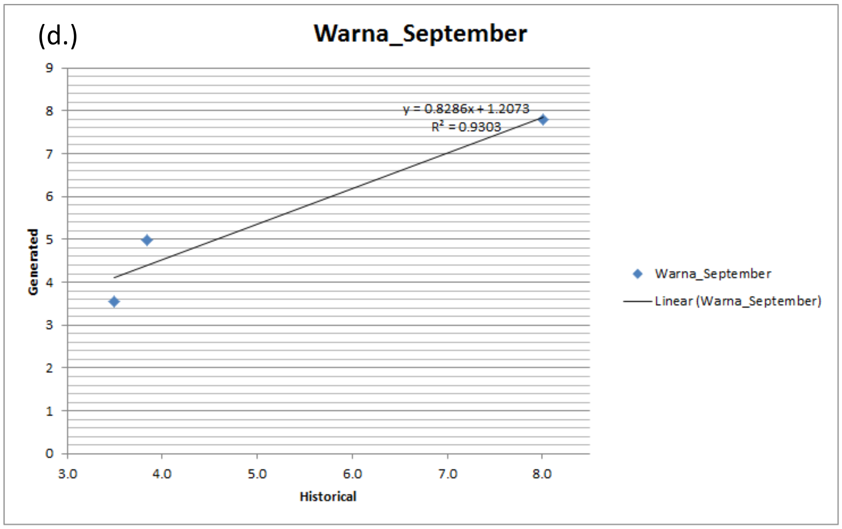


Figure 13: Plot of Historical Vs Generated Data for the station Warna

(a.)June (b.)July (c.)August (d.)September

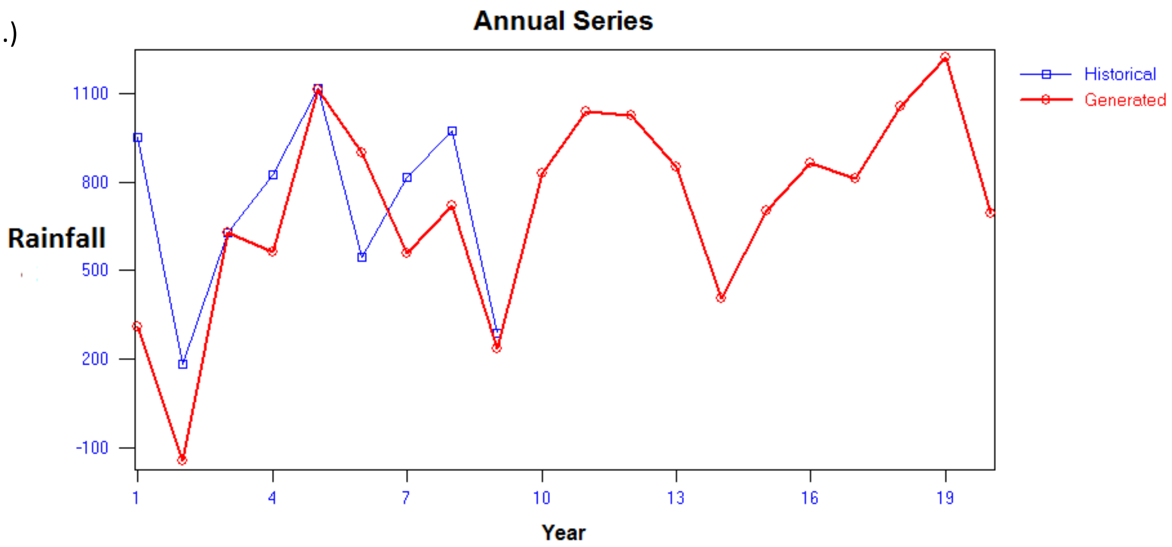
A6. Disaggregation by VS Model- Statistical Parameters

Table 8: Statistics of historical and generated data

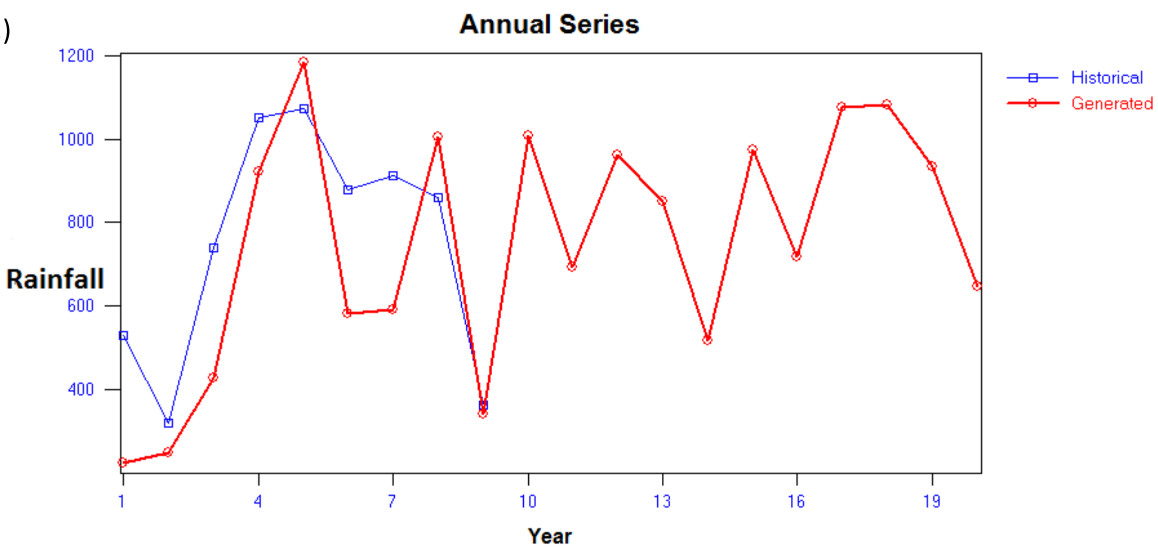
	Nal Lake		Sanand		Warna	
Parameter	**Historical**	**Generated**	**Historical**	**Generated**	**Historical**	**Generated**
Mean	703.4	666.3	747.7	706.2	414.9	392.1
Std. Dev	299.7	300.3	265.4	272.5	171.5	168.6
CV	0.4261	0.4472	0.3550	0.3817	0.4135	0.4251
Skew	-0.4346	0.1075	-0.4451	-0.02046	0.1425	0.1526
Min	183.5	114.3	319.4	182.8	184.0	93.34
Max	1120	1259	1073	1217	701	728.3
ACF (1)	-0.2646	-0.1194	0.3671	-0.04652	-0.1955	-0.1137
ACF (2)	-0.2183	-0.0237	-0.1362	-0.0458	-0.1293	0.02347

A7. Disaggregation by VS Model- Time Series plots of stations

(a.)



(b.)



(c.)

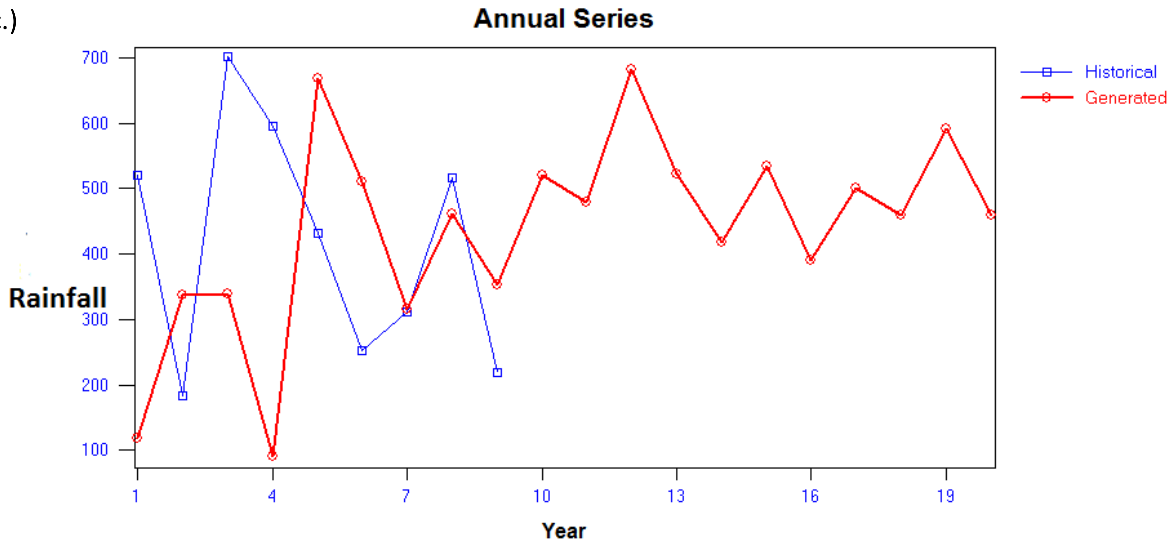


Figure 14: Plot of Historical Vs. Generated Data for a.) Nal Lake b.) Sanand c.) Warna